Use case: leukemia signatures (Den Boer, 2009)

Analyse Statistique des données Génomiques (ASG)

Jacques van Helden

2020-03-06

Use case: leukemia signatures (Den Boer, 2009)

Study case

We present here a dataset that will be used as study case to apply different approaches of multivariate analysis:

- data exploration
- multidimensional scaling
- differential analysis
- clustering (unsupervised classification)
- supervised classification

Reference: Den Boer ML *et al.* (2009). A subtype of childhood acute lymphoblastic leukaemia with poor treatment outcome: a genome-wide classification study. Lancet Oncol. 2009 10:125-34. [doi: 10.1016/S1470-2045(08)70339-5], [PMID 19138562]. Data available at Gene Expression Omnibus, series [GSE13425]

Identification of signatures for subtypes of leukemia

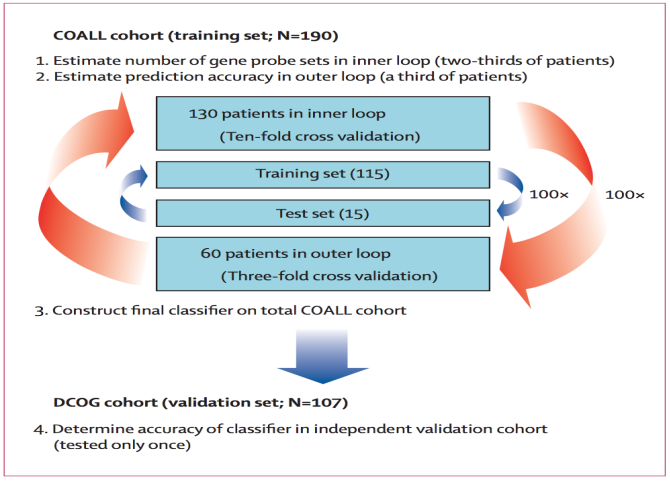

Figure 1: **Identification of a gene-expression signature enabling classification of paediatric ALL**

Figure 1: Figure 1 from Den Boer et al. (2009).

Hierarchical clustering of genes and samples

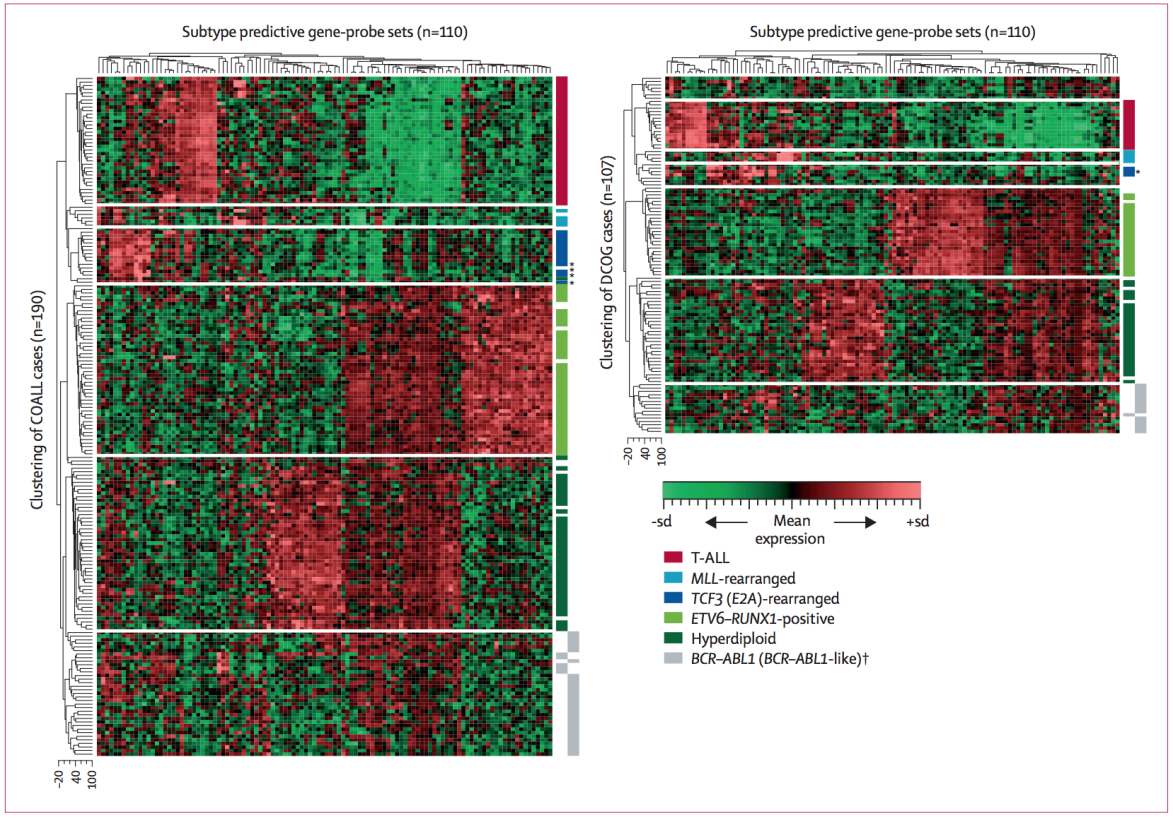

***Figure 2:* Clustering of ALL subtypes by gene-expression profiles**
Hierarchical clustering of patients from the COALL (left) and DCOG (right) studies with 110 gene-probe sets selected to classify paediatric ALL. Heat map shows which gene-probe sets are overexpressed (in red) and which gene probe sets are underexpressed (in green) relative to mean expression of all gene-probe sets (see scale bar). *Patients with *E2A*-rearranged subclone (15–26% positive cells). †Right column of grey bar denotes *BCR–ABL1*-like cases.

Figure 2: Figure 2 from Den Boer et al. (2009).

Survival curves

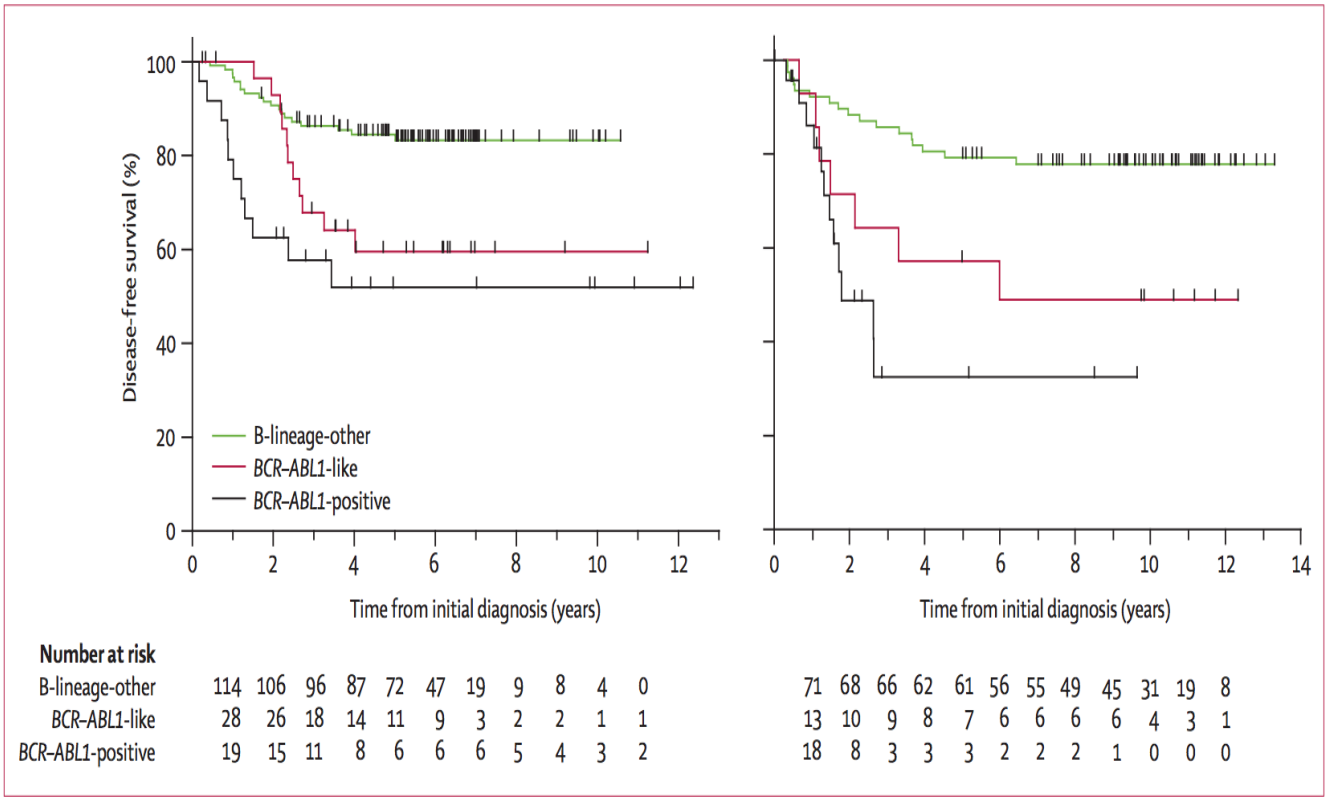

***Figure 3*: Kaplan-Meier estimates for the probability of disease-free survival (pDFS) in children with precursor B-ALL**
The COALL precursor B-ALL cohort (left) comprised 145 patients from COALL-92/97 and nine from DCOG ALL-9 treated at Sophia Children's Hospital. As reference, data for 22 patients with *BCR-ABL1*-positive disease enrolled in the COALL-92/97 protocol were included. Univariate analysis of pDFS comparing 30 *BCR-ABL1*-like with 119 remaining precursor B-ALL cases (excluding five *BCR-ABL1*-positives): p=0·012. Curves were almost the same when the nine DCOG patients were excluded. The DCOG precursor B-ALL cohort (right) comprised 92 children from DCOG-ALL8. As reference, data for 25 *BCR-ABL1*-positive cases enrolled in DCOG-ALL-7, 8, and 9 were included. Univariate analysis of pDFS comparing 14 *BCR-ABL1*-like and 77 remaining precursor B-ALL cases (excluding *BCR-ABL1*-positive case): p=0·026.

Figure 3: Figure 3 from Den Boer et al. (2009).